(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 2 936 060 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
11.10.2023 Bulletin 2023/41

(21) Application number: **13811098.6**

(22) Date of filing: **06.12.2013**

(51) International Patent Classification (IPC):
G01C 21/16 (2006.01) ***G01C 25/00*** (2006.01)
G06T 7/00 (2017.01) ***G06F 3/01*** (2006.01)
G06F 3/03 (2006.01) ***G06F 3/0346*** (2013.01)
G06T 7/80 (2017.01) ***G06T 7/32*** (2017.01)
G06F 3/0487 (2013.01) ***G06T 7/73*** (2017.01)

(52) Cooperative Patent Classification (CPC):
G06F 3/017; G01C 21/1656; G01C 25/005; G06F 3/0304; G06F 3/0346; G06F 3/0487; G06T 7/73; G06T 7/80; G06T 2207/30208

(86) International application number:
PCT/US2013/073511

(87) International publication number:
WO 2014/099410 (26.06.2014 Gazette 2014/26)

(54) **DISPLAY OF SEPARATE COMPUTER VISION BASED POSE AND INERTIAL SENSOR BASED POSE**

ANZEIGE VON SEPARATER COMPUTERSICHTBASIERTER HALTUNG SOWIE AUF INERTIALSENSORBASIERTER HALTUNG

AFFICHAGE DE POSE À BASE DE VISION ARTIFICIELLE ET DE POSE À BASE DE CAPTEUR INERTIEL SÉPARÉES

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priority: **21.12.2012 US 201213725175**

(43) Date of publication of application:
28.10.2015 Bulletin 2015/44

(73) Proprietor: **Qualcomm Incorporated**
San Diego, CA 92121-1714 (US)

(72) Inventors:
- **RAMACHANDRAN, Mahesh**
  San Diego, California 92121-1714 (US)
- **BRUNNER, Christopher**
  San Diego, California 92121-1714 (US)
- **RAMANANDAN, Arvind**
  San Diego, California 92121-1714 (US)
- **DIAZ SPINDOLA, Serafin**
  San Diego, California 92121-1714 (US)
- **CHARI, Murali Ramaswamy**
  San Diego, California 92121-1714 (US)

(74) Representative: **Bardehle Pagenberg Partnerschaft mbB**
Patentanwälte Rechtsanwälte
Prinzregentenplatz 7
81675 München (DE)

(56) References cited:
EP-A2- 1 501 051 **EP-A2- 1 708 139**

Description

BACKGROUND

Background Field

[0001] Embodiments of the subject matter described herein are related generally to using a computer vision based pose and an inertial sensor based pose separately in a visualization application, and more specifically to displaying separate graphics for the computer vision based pose and the inertial sensor based pose to prompt user movement of the mobile device.

Relevant Background

[0002] An increasing number of mobile devices, such as smartphones, tablet computers, cameras, etc., include various types of inertial sensors in addition to a camera. Inertial sensors, such as accelerometers and gyroscopes, are useful for tracking movement of a mobile device due to their ability to measure acceleration and angular velocity at high sampling rates. Many applications, such as augmented reality, indoor navigation, etc., use a combination of inertial sensors and computer vision for a determination of the pose (position and orientation) and tracking of the mobile device,.

[0003] In order for inertial sensors to be used for accurate position determination, particularly when combined with computer vision techniques, accurate calibration of the inertial sensors is required. Calibration of the inertial sensors includes alignment of inertial sensors with respect to camera, as well as calibration of scale and non-orthogonality parameters of the inertial sensors. Calibration of inertial sensors, however, is currently a time consuming factory process, which is required for each individual mobile device. During factory calibration of inertial sensors, a series of precise movements of the mobile device is performed and correlated to the response of the inertial sensors. While some approaches for calibration of sensors with user motion exist, the performance of such approaches is heavily dependent on supplying user motions that are similar in quality to factory motions. Instructing users to provide complicated and yet highly accurate motion is difficult. Without these rich motions, however, factory calibration, with its accompanying increase in cost, is the only way to get reliably obtain accurate calibration parameters.

[0004] EP 1708139 A2 describes a calibration method and apparatus. An instruction is detected which indicates that a physical object having a fixed plane where the physical object can be moved, and a virtual object rendered so as to be positioned on the plane where the physical object can be moved, are observed as being in a predetermined positional relationship. Then, a difference between a position and orientation value of the virtual object and a position and orientation value of the physical object is acquired at the time when the instruction is detected, so that the difference is used as a correction value to correct the positional difference between the physical object and the virtual object. This involves using a position and orientation sensor of the physical object.

[0005] EP 1501051 A2 discloses a position and orientation detection method and apparatus. A method and an apparatus are provided for inputting a signal indicating a change in position of an image sensing apparatus to predict present position and orientation of the image sensing apparatus using previously obtained position and orientation, by identifying corresponding pairs of indicators detected from an image in real space sensed by the image sensing apparatus and indicators contained in an image when sensed from a predicted position and orientation and using a correction method determined by the number of identified indicator pairs either to correct all or some of the degrees of freedom of the predicted camera position and orientation or to use the predicted camera position and orientation directly as is. As a result, highly accurate detection results can be obtained using a simple structure in the detection of the position and orientation of the image sensing apparatus.

[0006] There is still a need for a more efficient way of increasing accuracy of determining a pose of a mobile device.

SUMMARY

[0007] The scope of protection of the present invention is defined by the independent claims.

[0008] Optional variants are defined in the dependent claims.

[0009] In one implementation, a method is provided, as defined by claim 1.

[0010] In one implementation, a mobile device is provided, as defined by claim 4.

[0011] In one implementation, a computer program is provided, as defined by claim 8.

[0012] In one implementation, a storage medium including program code stored thereon is provided, as defined by claim 7.

BRIEF DESCRIPTION OF THE DRAWING

[0013]

Fig. 1 illustrates a mobile device capable of determining a vision based pose from captured images and a sensor based pose from inertial sensors and using the vision based pose and the sensor based pose separately in a visualization application.

Fig. 2 illustrates one possible implementation of graphics that may be used by the mobile device in the visualization application.

Fig. 3 illustrates another possible implementation of graphics that may be used by the mobile device in the visualization application.

Fig. 4 illustrates another possible implementation of graphics that may by the mobile device in the visualization application.

Fig. 5 is a flow chart illustrating a process of using a vision based graphic and a sensor based graphic separately in a visualization application.

Fig. 6 is a flow chart illustrating a process of using a vision based graphic, a sensor based graphic, and a hybrid graphic in a visualization application.

Fig. 7 is a flow chart illustrating a specific embodiment of the visualization application for calibrating the inertial sensors.

Fig. 8 is a block diagram of a mobile device capable of determining a vision based pose and a sensor based pose and using a vision based graphic and a sensor based graphic separately in a visualization application.

DETAILED DESCRIPTION

[0014] Fig. 1 illustrates a mobile device 100 capable of determining a vision based pose from captured images and a sensor based pose from inertial sensors and using the vision based pose and the sensor based pose separately in a visualization application. The visualization application displays separate graphics for the vision based pose and the sensor based pose, sometimes respectively referred to herein as a vision based graphic and a sensor based graphic. The visualization application, by way of example, may be used to calibrate the inertial sensors, gaming, photography, etc. The inertial sensors may be calibrated, for example, based on user manipulation of the mobile device in response to the displayed graphics. The mobile device 100 is illustrated as including a display 102, which may be, but is not necessarily, a touch screen display. The mobile device 100 includes a forward facing camera 108 to image the environment, such as the target 101, which is illustrated as being shown on display 102. The mobile device 100 also includes inertial sensors 110, such as accelerometers, gyroscopes or the like. The mobile device 100 may also include other features that are not relevant to the present disclosure, such as a speaker 104 and microphone 106, e.g., if the mobile device 100 is a cellular telephone.

[0015] As used herein, a "mobile device" refers to any portable electronic device capable of vision-based position detection and tracking from captured images or video streams and that includes inertial sensors, and may include a e.g., cellular or other wireless communication device, personal communication system (PCS) device, personal navigation device (PND), Personal Information Manager (PIM), Personal Digital Assistant (PDA), or other suitable mobile device including cameras, wireless communication devices, computers, laptops, tablet computers, etc. The mobile device may be, but need not necessarily be capable of receiving wireless communication and/or navigation signals, such as navigation positioning signals. The term "mobile device" is also intended to include devices which communicate with a personal navigation device (PND), such as by short-range wireless, infrared, wireline connection, or other connection - regardless of whether satellite signal reception, assistance data reception, and/or position-related processing occurs at the device or at the PND.

[0016] The mobile device 100 includes a vision pose module 112 that uses the images captured by the camera 108, e.g., the series of video frames produced by the camera 108, to generate a vision based pose. A vision based pose is a position and orientation of the mobile device that is determined using computer vision techniques. It should be understood that the camera 108 may capture images and/or frames of video, and that the terms image and frame are used interchangeably herein. The mobile device 100 further includes an inertial sensory system (INS) 114 that uses the output of the inertial sensors 110 to generate the sensor based pose. It should be understood that one or both of the vision based pose and the sensor based pose may be initialized based on the other, but otherwise, as used herein, the vision based pose is independent of the sensor based pose, and the sensor based pose is independent of the vision based pose. Thus, at a current time instant the vision based pose and the sensor based pose are independent when rendering a current set of augmentations; but they could have been fused together at a past time instant.

[0017] The mobile device 100 uses a visualization application that displays separate graphics on the display 102 based on the vision based pose generated from the captured images and the sensor based pose generated from inertial sensor data. For example, as illustrated in Fig. 1, the mobile device 100 may render a first graphic 120 based on the vision based pose generated by the vision pose module 112 and a second graphic 122 based on the sensor based pose generated by the INS 114. If desired, the vision based graphic 120 and the sensor based graphic 122 may be different color channels of the same rendered object or they may be different objects. Fig. 1 illustrates the graphics 120 and 122 as balls, but it should be understood that any type of graphics may be used. Moreover, if desired, multiple graphics may be used for the vision based pose and multiple graphics may be used for the sensor based pose. For example, as discussed below, one or more horizontal bars and one or more vertical bars may be rendered based on the vision based pose and, similarly, horizontal and vertical bars may be rendered based on the sensor based pose. If desired, the rendered graphics may be more complicated or animated, e.g., a cat may be rendered based on the vision based pose and a mouse rendered based on the sensor based pose, wherein the cat is controlled to catch the mouse based on user motion.

[0018] The visualization application separately displays the vision based graphic and sensor based graphic to prompt user movement of the mobile device. For example, the visualization application may be entertainment, e.g., a game. Another application may be related to photography. For example, to produce a three-dimensional image using a single camera, the mobile device is moved to different perspectives, while maintaining the imaged subject centered in the display. Thus, a vision based graphic may be displayed stationary relative to the subject, while movement of the sensor based graphic, prompts the user to move the mobile device to different perspectives in order to maintain alignment of the vision based graphic and sensor based graphic. In another example, to produce a panorama, it is desirable to maintain the position of the camera while panning the camera. Thus, a sensor based graphic may be displayed as stationary in the display, while movement of the vision based graphic, prompts the user to rotate the mobile device while holding the position stationary in order to maintain alignment of the vision based graphic and sensor based graphic.

[0019] The visualization application is to be used for calibration of the inertial sensors 110. Calibration of the alignment as well as scale and non-orthogonality of the inertial sensors 110 and the camera 108 may use the vision based poses as external measurements. The user is prompted to move the mobile device 100 with desired trajectories for the calibration based on the display of the vision based graphic and sensor based graphic. Thus, the user is incentivized move the mobile device along desired trajectories to collect good-quality data which can be fed into the calibration algorithm.

[0020] The mobile device 100 collects data from the inertial sensors 110 produced in response to movement of the mobile device 100 while capturing images that are used to generate the vision based pose. The vision based pose and sensor data can be used for calibration of the inertial sensors. The vision based graphic 120 and sensor based graphic 122 are rendered to prompt the user to move the mobile device, as well as feedback to the user with respect to the status of the calibration of the inertial sensors, such as the calibration accuracy and confidence in calibration parameter. Thus, the mobile device 100 may use the visualization application to allow the user to visualize the inertial sensor calibration accuracy and adaptively change user interaction with the mobile device based on the confidence in the calibration parameter estimates.

[0021] Fig. 2, by way of example, illustrates one possible implementation of graphics that may be used by the mobile device in the visualization application when calibrating the inertial sensors. As can be seen in comparison to the pose of the mobile device 100 shown in Fig. 1, the mobile device 100 in Fig. 2 is illustrated as rotated with respect to the target 101. Due to the rotation of the mobile device 100, the sensor based graphic 122 has moved to closer alignment with the vision based graphic 120. The process of manually moving the mobile device 100 to align the graphics may provide feedback to calibrate the inertial sensors 110 with respect to the camera 108. Moreover, the visual display of the separate graphics 120 and 122 may provide feedback to the user of the status of the calibration of the inertial sensors 110 to the camera 108. Additionally, as illustrated by text 124 in Fig. 2, a score may be provided to the user to indicate the status of the alignment. Other or additional feedback may be provided to the user as well, such as the horizontal bar 126H and vertical bar 126V, which may indicate the status of alignment along the horizontal and vertical axes, respectively. Moreover, as there are six dimensions, three for rotation and three for translation, six separate graphics, e.g., bars or otherwise, or a subset therefore, e.g., two or more, may be displayed. The graphics 120 and 122 may be rendered as part of a game that encourages the user to move the mobile device in a specific direction to accelerate the convergence of the calibration of the inertial sensors. In other words, the calibration procedure may be presented to the user as a game to be played for entertainment, as opposed to a burdensome administrative process. For example, in a calibration game, the displayed position of the vision based graphic 120 may be periodically moved requiring the user to manipulate the position and orientation of the mobile device 100 to re-align the sensor based graphic 122 with the vision based graphic 120. Thus, the user may be prompted to rotate and/or move the mobile device 100 along one or more axes, and the resulting motion of the mobile device 100 may be used for calibration. With increasingly better calibration, the motion estimated by the inertial sensors will match increasingly closer to the motion solved based on computer vision. The calibration may be determined to be acceptable when the difference between the motion estimated by the inertial sensors and the motion solved based on computer vision matches within a threshold, which may be based on the noise level of the system.

[0022] Alignment and non-orthogonality are expected to change very slowly over the life time of the phone. Hence, calibrating alignment and non-orthogonality may be performed infrequently. Scale, however, depends on temperature. While scale can be estimated with bias in an on-going manner, jointly estimating alignment, non-orthogonality, scale, and bias will likely lead to decreased accuracy in the overall calibration parameters. Accordingly, user calibration of the mobile device 100 may be periodically performed or in the event of mechanical stresses on the mobile device 100.

[0023] Fig. 3 illustrates another possible implementation of graphics that may be used by the mobile device in the visualization application when calibrating the inertial sensors. A displayed graphic 132 is based on the computer vision pose and a separate graphic 134 may be based on the sensor based pose. As illustrated in Fig. 3, the vision based graphic 132 may be a maze while the sensor based graphic 134 may be an object to be moved through the maze by user generated motion of the mobile device 100. Thus, as can be seen in Fig. 3, the user will be prompted to rotate or move the mobile device 100 horizontally and vertically to move the sensor based graphic 134 through the vision based graphic 132. The motion of the mobile device 100 is used to calibrate the inertial sensors 110.

[0024] Fig. 4 illustrates another possible implementation of graphics that may be used by the mobile device in the visualization application when calibrating the inertial sensors. A displayed graphic 142 is based on the computer vision pose and a separate graphic 144 is based on the sensor based pose. The vision based graphic 142 and the sensor based graphic 144 may be used, for example, to prompt the user to rotate the mobile device 100, e.g., along the Z axis extending normal to the display 102. Additionally, the user may prompted to move the mobile device 100 forwards and back, e.g., translate the mobile device along the Z axis along the axis that extends normal to the display 102, to alter the size of one or more of the graphics 142, 144 until the graphics are similar in size. Similarly, the user may be prompted to move (translate) the mobile device 100 along the X and Y axes to align the graphics 142, 144 and may be prompted to rotate the mobile device 100 about the X and Y axes based on the appearance of the graphics 142, 144. Thus, it can be seen that the user may be prompted to move the mobile device with 6 degrees of freedom.

[0025] Fig. 5 is a flow chart illustrating a process of using a vision based graphic and a sensor based graphic separately in a visualization application. A vision based pose of the mobile device is determined using captured images (202). The vision based pose of the mobile device may be produced using conventional computer vision techniques. For example, for each image new image, features may be extracted using FAST (Features from Accelerated Segment Test) corner detector, Scale Invariant Feature Transform (SIFT), Speeded-up Robust Features (SURF). The extracted features for the current image may be compared to features from a reference image, which may be a previous image or model of a known or learned target. A homography mapping the positions of the extracted features in the current image to the reference image may provide a pose of the mobile device with respect to the environment with 6 degrees of freedom.

[0026] A sensor based pose of the mobile device using inertial sensors is also determined (204). As is well known, a sensor based pose may be determined by integrating the inertial sensor data, e.g., from accelerometers or gyroscopes, over time to determine a change in the pose of the mobile device over the integration time. Thus, by initiating integration of the inertial sensors at the same time as capturing a reference image of the environment with the camera, the sensor based pose of the mobile device can be determined in parallel with determining the vision based pose. It should be understood that the data from the inertial sensors provides a relative change in the pose. The change in pose may be used as the sensor based pose. Alternatively, an initial pose may be provided, e.g., using the computer vision technique or other initialization technique, where the sensor based pose may be deterred as a change from the initial pose. For example, the sensor based pose may be initialized with the vision based pose at regular intervals, e.g., every 5th frame. Between those initialization frames, the change in pose from sensors is used as sensor based pose.

[0027] The vision based pose and the sensor based pose are used separately in a visualization application (206). In other words, the vision based pose and the sensor based pose are used at the same time by the same visualization application, but are not combined so that they are used separately. For example, in the visualization application, a first graphic based on the vision based pose and a second graphic based on the sensor based pose are displayed. If desired, the vision based pose and the sensor based pose may be used in the visualization application by rendering a single augmentation character with different color channels. Thus, the first graphic may be a first color channel for the single augmentation character and the second graphic may be a second color channel for the single augmentation character.

[0028] As discussed above, the visualization application performs calibration of the inertial sensors using the captured images, wherein the first graphic and the second graphic may be rendered to interact with each other based on the calibration of the inertial sensors. The user is prompted to move the mobile device in a specific direction, for example by moving at least one of the first graphic and second graphic based on a status of a calibration of the inertial sensors, wherein movement of the mobile device in the specific direction accelerates convergence of the calibration of the inertial sensors. For example, after completion of calibration of one calibration parameter, e.g., the X axis, the first graphic and/or the second graphic may be moved to prompt the user to move the mobile device along the Y axis. The Z axis may be similarly calibrated, as well as rotations about the various axes. Additionally, or alternatively, the status of the calibration of the inertial sensors may be, e.g., the calibration accuracy, a confidence in a calibration parameter, an error covariance or a variation of the error covariance. An example of the error covariance or confidence is the result from an Extended Kalman Filter, which may be used for the calibration procedure. Another example could be from another algorithm that estimates the reliability of the estimates. Prompting the user to move the mobile device in the specific direction uses the quality of the calibration of the inertial sensors as feedback to the visualization application to accelerate the convergence of the calibration of the inertial sensors.

[0029] If desired, additional information may be displayed by the visualization application in the mobile device 100. For example, Fig. 6 is a flow chart, similar to that shown in Fig. 5, but includes additional information used by the visualization application. Thus, as can be seen, the vision based pose and senor based pose are determined (202, 204). Additionally, a hybrid based pose of the mobile device is determined using the captured images and the inertial sensors together (212). For example, if desired, the vision based pose and the sensor based pose may be combined, or the data from the inertial sensors and data from the camera may be combined to form the hybrid based pose. The visualization application uses the vision based pose, sensor based pose, and the hybrid based pose separately (214), e.g., by displaying separate graphics representing each pose. It should be understood that one or both of the graphics displayed based on the vision based pose or the sensor based pose need not be displayed at all times during calibration, e.g., if tracking of the vision based pose is lost, the graphic displayed for the vision based pose may be removed, while the graphic displayed for the sensor based pose is still displayed.

[0030] Fig. 7 is a flow chart illustrating a specific embodiment of the visualization application in the form of a calibration application for the inertial sensors. As can be seen, the calibration process is initiated (302). The calibration application may be in place of or in addition to the factory calibration and may be pre-installed on the mobile device 100. It may be desirable to automatically initiate the calibration application prior to permitting the use of inertial sensors or, alternatively, the user may initiate the application when desired. The user may be provided with instructions, such as information as to where to find a feature rich target that is appropriate for producing accurate computer vision pose estimates. The mobile device 100 may be placed in a specific position, e.g., face up on a level surface, for initializing the pose of the mobile device (304). The mobile device 100 may indicate to the user when initializing is complete, e.g., by emitting a tone. The mobile device 100 determines a sensor based pose and a separate vision based pose (306). The mobile device 100 captures images or video frames of a reference target to determine a pose and acquire tracking using computer vision, while also determining a pose based on inertial sensors, which are also used for tracking. Instructions are provided to the user for moving the mobile device by displaying separate graphics related to the sensor based pose and vision based pose (308). The sensor based pose and vision based pose are individually tracked while the user moves the mobile device to improve calibration (310).

[0031] The instructions for moving the mobile device may be generated based on the confidence of each pose estimate, which may be derived from the error covariance from Extended Kalman Filters or from other algorithms. For instance, if the confidence is low for a pose generated by a gyroscope measuring rotation around the Z axis, the user may be instructed to rotate the mobile device around the Z axis while continuing to capture images of the reference target until a threshold value for the confidence is reached. The confidence threshold may be expressed in terms of a score or changes in the displayed graphics to encourage a user to make specific motions to increase the confidence of estimates. Providing a score and/or challenges to the user leverages on a prevalent behavior of a game player to play in order to increase his/her score in a game or finish a challenge. Each time the calibration of one of the calibration parameters is complete, i.e., the confidence level increases beyond a threshold, the user is awarded points. The time it takes the user to complete may also be included in the score. Thus, users are incentivized to manipulate the mobile device to provide input data for the calibration process. Accordingly, user interaction with the mobile device may be used for calibration, thereby obviating the need for expensive factory calibration of the inertial sensors.

[0032] Fig. 8 is a block diagram of a mobile device 100 capable of determining a vision based pose and a sensor based pose and using a vision based graphic and a sensor based graphic separately in a visualization application. The mobile device 100 includes a camera 108 as well as inertial sensors 110, such as accelerometers, gyroscopes or the like. The mobile device 100 further includes a user interface 150, which includes the display 102, as well as a keypad 152 or other input device through which the user can input information into the mobile device 100. If desired, the keypad 152 may be obviated by integrating a virtual keypad into the display 102 with a touch sensor (or gesture control). The user interface 150 may also include the microphone 106 and speaker 104, e.g., if the mobile device 100 is a cellular telephone or the like. Of course, mobile device 100 may include other elements unrelated to the present disclosure.

[0033] The mobile device 100 also includes a control unit 105 that is connected to and communicates with the camera 108, inertial sensors 110, as well as the user interface 150, including the display 102. The control unit 105 may be provided by a bus 105b, processor 105p and associated memory 105m, hardware 105h, firmware 105f, and software 105s. The control unit 105 receives and processes images captured by the camera as well as data obtained from the inertial sensors 110, as discussed above. The control unit 105 is further illustrated as including a vision pose module 112 that uses the images captured by the camera 108 to generate the vision based pose using computer vision techniques. The mobile device 100 further includes the INS 114 that uses the output of the inertial sensors 110 to generate the sensor based pose. A rendering module 116 generates a vision based graphic and a sensor based graphic and any other graphical data to be provided to the user, which is shown on display 102. Additionally, a visualization application module 118 uses the vision based pose and sensor based pose and causes the graphics from rendering module 116 to be displayed. For example, the visualization application module 118 may be a calibration module that calibrates the inertial sensors 110 based on the vision based pose and the sensor based pose, as discussed above, or may alternatively be a motion based game or other appropriate type of application, such as a photography application. A hybrid pose module 119 may be included to determine a pose of the mobile device based on the combined vision based pose and sensor based pose, or otherwise determined based on the captured images and inertial sensor data.

[0034] The vision pose module 112, INS 114, rendering module 116, visualization application module 118, and hybrid pose module 119 are illustrated separately from processor 105p for clarity, but may be part of the processor 105p or implemented in the processor based on instructions in the software 105s which is run in the processor 105p. It will be understood as used herein that the processor 105p can, but need not necessarily include, one or more microprocessors, embedded processors, controllers, application specific integrated circuits (ASICs), digital signal processors (DSPs), and the like. The term processor is intended to describe the functions implemented by the system rather than specific hardware. Moreover, as used herein the term "memory" refers to any type of computer storage medium, including long term, short term, or other memory associated with the mobile device, and is not to be limited to any particular type of memory or number of memories, or type of media upon which memory is stored.

[0035] The methodologies described herein may be implemented by various means depending upon the application. For example, these methodologies may be implemented in hardware 105h, firmware 105f, software 105s, or any combination thereof. For a hardware implementation, the processing units may be implemented within one or more application specific integrated circuits (ASICs), digital signal processors (DSPs), digital signal processing devices (DSPDs), programmable logic devices (PLDs), field programmable gate arrays (FPGAs), processors, controllers, micro-controllers, microprocessors, electronic devices, other electronic units designed to perform the functions described herein, or a combination thereof.

[0036] For a firmware and/or software implementation, the methodologies may be implemented with modules (e.g., procedures, functions, and so on) that perform the functions described herein. Any machine-readable medium tangibly embodying instructions may be used in implementing the methodologies described herein. For example, software codes may be stored in memory 105m and executed by the processor 105p. Memory 105m may be implemented within or external to the processor 105p. If implemented in firmware and/or software, the functions may be stored as one or more instructions or code on a storage medium that is computer-readable, wherein the storage medium does not include transitory propagating signals. Examples include storage media encoded with a data structure and storage encoded with a computer program. Storage media includes physical computer storage media. A storage medium may be any available medium that can be accessed by a computer. By way of example, and not limitation, such storage media can comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other medium that can be used to store desired program code in the form of instructions or data structures and that can be accessed by a computer; disk and disc, as used herein, includes compact disc (CD), laser disc, optical disc, digital versatile disc (DVD), floppy disk and Blu-ray disc where disks usually reproduce data magnetically, while discs reproduce data optically with lasers. Combinations of the above should also be included within the scope of storage media.

[0037] Although the present invention is illustrated in connection with specific embodiments for instructional purposes, the present invention is not limited thereto. Various adaptations and modifications may be made without departing from the scope of the appended claims.

Claims

1. A method performed by a mobile device (100), the method comprising:

    determining a vision based pose of the mobile device (100) using images captured by a camera (108) of the mobile device (100) by extracting one or more features and comparing the extracted features to features from a reference image, wherein the reference image is a previous image or model of a known or learned target;
    determining a sensor based pose of the mobile device (100) using inertial sensors (110) of the mobile device (100), wherein the vision based pose is a first estimate of a position and orientation of the mobile device (100) that is determined using computer vision techniques and the sensor based pose is a second estimate of the position and orientation of the mobile device (100) that is determined using inertial sensors (110) of the mobile device (100); and
    using the vision based pose and the sensor based pose separately in a visualization application, wherein using the vision based pose and the sensor based pose separately in the visualization application comprises displaying a first graphic based on the vision based pose and a second graphic based on the sensor based pose;
    wherein using the vision based pose and the sensor based pose separately in the visualization application comprises performing calibration of the inertial sensors (no), wherein a user is prompted to move the mobile device (100) along trajectories for the calibration based on the first graphic and the second graphic to collect

data which are fed into a calibration algorithm.

2. The method of claim 1, further comprising determining a hybrid based pose of the mobile device (100) based on combining the vision based pose and the sensor based pose or on combining data from the inertial sensors (110) and data from the camera (108) and using the hybrid based pose in the visualization application.

3. The method of claim 1, wherein using the vision based pose and the sensor based pose separately in the visualization application comprises: prompting a user to move the mobile device (100) in a specific direction with the first graphic and the second graphic, wherein movement of the mobile device (100) in the specific direction accelerates convergence of the calibration of the inertial sensors (no).

4. A mobile device (100) comprising:

   means for determining a vision based pose of the mobile device (100) using images captured by a camera (108) of the mobile device (100) by extracting one or more features and comparing the extracted features to features from a reference image, wherein the reference image is a previous image or model of a known or learned target;
   means for determining a sensor based pose of the mobile device (100) using inertial sensors (110) of the mobile device (100), wherein the vision based pose is a first estimate of a position and orientation of the mobile device (100) that is determined using computer vision techniques and the sensor based pose is a second estimate of the position and orientation of the mobile device (100) that is determined using inertial sensors (110) of the mobile device (100); and
   means for using the vision based pose and the sensor based pose separately in a visualization application, wherein the means for using the vision based pose and the sensor based pose separately in the visualization application displays a first graphic based on the vision based pose and a second graphic based on the sensor based pose,
   wherein using the vision based pose and the sensor based pose separately in the visualization application comprises performing calibration of the inertial sensors (no), wherein a user is prompted to move the mobile device (100) along trajectories for the calibration based on the first graphic and the second graphic to collect data which are fed into a calibration algorithm.

5. The mobile device (100) of claim 4, further comprising means for determining a hybrid based pose of the mobile device (100) based on combining the vision based pose and the sensor based pose or on combining data from the inertial sensors (110) and data from the camera (108) and wherein the means for using the vision based pose and the sensor based pose separately in the visualization application further uses the hybrid based pose.

6. The mobile device (100) of claim 4, wherein the means for using the vision based pose and the sensor based pose separately in the visualization application prompts a user to move the mobile device (100) in a specific direction with the first graphic and the second graphic, wherein movement of the mobile device (100) in the specific direction accelerates convergence of the calibration of the inertial sensors (110).

7. A storage medium including program code stored thereon, comprising: program code that, when executed on a processor of a mobile device (100) equipped with a camera (108) and one or more inertial sensors (110), causes the processor to perform the method of any one of claims 1 to 3.

8. A computer program comprising program code that, when executed on a processor of a mobile device (100) equipped with a camera (108) and one or more inertial sensors (110), causes the processor to perform the method of any one of claims 1 to 3.

Patentansprüche

1. Verfahren, das von einer mobilen Vorrichtung (100) durchgeführt wird, wobei das Verfahren umfasst:

   Bestimmen einer sichtbasierten Stellung der mobilen Vorrichtung (100) unter Verwendung von Bildern, die von einer Kamera (108) der mobilen Vorrichtung (100) aufgenommen werden, durch Extrahieren eines oder mehrerer Merkmale und Vergleichen der extrahierten Merkmale mit Merkmalen aus einem Referenzbild, wobei das Referenzbild ein vorheriges Bild oder Modell eines bekannten oder erlernten Ziels ist;
   Bestimmen einer sensorbasierten Stellung der mobilen Vorrichtung (100) unter Verwendung von Trägheitssensoren (110) der mobilen Vorrichtung (100), wobei die sichtbasierte Stellung eine erste Schätzung einer Position und Ausrichtung der mobilen Vorrichtung (100) ist, die unter Verwendung von Computervisionstechniken bestimmt wird, und die sensorbasierte Stellung eine zweite Schätzung der Position und Ausrichtung der mobilen Vorrichtung (100) ist, die unter Verwendung von Trägheitssensoren

(110) der mobilen Vorrichtung (100) bestimmt wird; und
getrenntes Verwenden der sichtbasierten Stellung und der sensorbasierten Stellung in einer Visualisierungsanwendung, wobei das getrennte Verwenden der sichtbasierten Stellung und der sensorbasierten Stellung in der Visualisierungsanwendung Anzeigen einer ersten Grafik basierend auf der sichtbasierten Stellung und einer zweiten Grafik basierend auf der sensorbasierten Stellung umfasst;
wobei das getrennte Verwenden der sichtbasierten Stellung und der sensorbasierten Stellung in der Visualisierungsanwendung Durchführen einer Kalibrierung der Trägheitssensoren (110) umfasst, wobei ein Benutzer aufgefordert wird, die mobile Vorrichtung (100) entlang von Trajektorien für die Kalibrierung basierend auf der ersten Grafik und der zweiten Grafik zu bewegen, um Daten zu sammeln, die in einen Kalibrierungsalgorithmus eingespeist werden.

2. Verfahren nach Anspruch 1, ferner umfassend Bestimmen einer hybridbasierten Stellung der mobilen Vorrichtung (100) basierend auf Kombinieren der sichtbasierten Stellung und der sensorbasierten Stellung oder auf Kombinieren von Daten von den Trägheitssensoren (110) und Daten von der Kamera (108) und Verwenden der hybridbasierten Stellung in der Visualisierungsanwendung.

3. Verfahren nach Anspruch 1, wobei das getrennte Verwenden der sichtbasierten Stellung und der sensorbasierten Stellung in der Visualisierungsanwendung umfasst:
Auffordern eines Benutzers, die mobile Vorrichtung (100) in einer spezifischen Richtung mit der ersten Grafik und der zweiten Grafik zu bewegen, wobei Bewegung der mobilen Vorrichtung (100) in der spezifischen Richtung Konvergenz der Kalibrierung der Trägheitssensoren (110) beschleunigt.

4. Mobile Vorrichtung (100), umfassend:

Mittel zum Bestimmen einer sichtbasierten Stellung der mobilen Vorrichtung (100) unter Verwendung von
Bildern, die von einer Kamera (108) der mobilen Vorrichtung (100) aufgenommen werden, durch Extrahieren eines oder mehrerer Merkmale und Vergleichen der extrahierten Merkmale mit Merkmalen aus einem Referenzbild, wobei das Referenzbild ein vorheriges Bild oder Modell eines bekannten oder erlernten Ziels ist;
Mittel zum Bestimmen einer sensorbasierten Stellung der mobilen Vorrichtung (100) unter Verwendung von Trägheitssensoren (110) der mobilen Vorrichtung (100), wobei die sichtbasierte Stellung eine erste Schätzung einer Position und Ausrichtung der mobilen Vorrichtung (100) ist, die unter Verwendung von Computervisionstechniken bestimmt wird, und die sensorbasierte Stellung eine zweite Schätzung der Position und Ausrichtung der mobilen Vorrichtung (100) ist, die unter Verwendung von Trägheitssensoren (110) der mobilen Vorrichtung (100) bestimmt wird; und
Mittel zum getrennten Verwenden der sichtbasierten Stellung und der sensorbasierten Stellung in einer Visualisierungsanwendung, wobei das Mittel zum getrennten Verwenden der sichtbasierten Stellung und der sensorbasierten Stellung in der Visualisierungsanwendung eine erste Grafik basierend auf der sichtbasierten Stellung und eine zweite Grafik basierend auf der sensorbasierten Stellung anzeigt,
wobei das getrennte Verwenden der sichtbasierten Stellung und der sensorbasierten Stellung in der Visualisierungsanwendung Durchführen einer Kalibrierung der Trägheitssensoren (110) umfasst, wobei ein Benutzer aufgefordert wird, die mobile Vorrichtung (100) entlang von Trajektorien für die Kalibrierung basierend auf der ersten Grafik und der zweiten Grafik zu bewegen, um Daten zu sammeln, die in einen Kalibrierungsalgorithmus eingespeist werden.

5. Mobile Vorrichtung (100) nach Anspruch 4, ferner umfassend Mittel zum Bestimmen einer hybridbasierten Stellung der mobilen Vorrichtung (100) basierend auf Kombinieren der sichtbasierten Stellung und der sensorbasierten Stellung oder auf Kombinieren von Daten von den Trägheitssensoren (110) und Daten von der Kamera (108) und wobei das Mittel zum getrennten Verwenden der sichtbasierten Stellung und der sensorbasierten Stellung in der Visualisierungsanwendung ferner die hybridbasierte Stellung verwendet.

6. Mobile Vorrichtung (100) nach Anspruch 4, wobei das Mittel zum getrennten Verwenden der sichtbasierten Stellung und der sensorbasierten Stellung in der Visualisierungsanwendung einen Benutzer auffordert, die mobile Vorrichtung (100) in einer spezifischen Richtung mit der ersten Grafik und der zweiten Grafik zu bewegen, wobei Bewegung der mobilen Vorrichtung (100) in der spezifischen Richtung Konvergenz der Kalibrierung der Trägheitssensoren (110) beschleunigt.

7. Speichermedium, das darauf gespeicherten Programmcode beinhaltet, umfassend:
Programmcode, der, wenn er auf einem Prozessor einer mobilen Vorrichtung (100) ausgeführt wird, die mit einer Kamera (108) und einem oder mehreren Trägheitssensoren (110) ausgestattet ist, den Pro-

zessor veranlasst, das Verfahren nach einem der Ansprüche 1 bis 3 durchzuführen.

8. Computerprogramm, umfassend Programmcode, der, wenn er auf einem Prozessor einer mobilen Vorrichtung (100) ausgeführt wird, die mit einer Kamera (108) und einem oder mehreren Trägheitssensoren (110) ausgestattet ist, den Prozessor veranlasst, das Verfahren nach einem der Ansprüche 1 bis 3 durchzuführen.

Revendications

1. Un procédé réalisé par un dispositif mobile (100), le procédé comprenant :

la détermination d'une pose du dispositif mobile (100) basée sur la vision, en utilisant des images capturées par une caméra (108) du dispositif mobile (100), en extrayant une ou plusieurs caractéristiques et en comparant les caractéristiques extraites à des caractéristiques provenant d'une image de référence, dans lequel l'image de référence est une image précédente ou un modèle d'une cible connue ou apprise ;
la détermination d'une pose du dispositif mobile (100) basée sur des capteurs, en utilisant des capteurs inertiels (110) du dispositif mobile (100), dans lequel la pose basée sur la vision est une première estimation d'une position et d'une orientation du dispositif mobile (100) qui est déterminée en utilisant des techniques de vision par calculateur, et
la pose basée sur des capteurs est une deuxième estimation de la position et de l'orientation du dispositif mobile (100) qui est déterminée en utilisant des capteurs inertiels (110) du dispositif mobile (100) ; et
l'utilisation séparée de la pose basée sur la vision et de la pose basée sur des capteurs dans une application de visualisation, dans lequel l'utilisation séparée de la pose basée sur la vision et de la pose basée sur des capteurs dans l'application de visualisation comprend l'affichage d'un premier graphisme basé sur la pose basée sur la vision et d'un deuxième graphisme basé sur la pose basée sur des capteurs ;
dans lequel l'utilisation séparée de la pose basée sur la vision et de la pose basée sur des capteurs dans l'application de visualisation comprend la réalisation d'un étalonnage des capteurs inertiels (110), dans lequel un utilisateur est invité à déplacer le dispositif mobile (100) le long de trajectoires pour l'étalonnage sur la base du premier graphisme et du deuxième graphisme pour collecter des données qui sont introduites dans un algorithme d'étalonnage.

2. Le procédé selon la revendication 1, comprenant en outre la détermination d'une pose du dispositif mobile (100) à base hybride sur la base d'une combinaison de la pose basée sur la vision et de la pose basée sur des capteurs, ou d'une combinaison de données provenant des capteurs inertiels (110) et de données provenant de la caméra (108), et l'utilisation de la pose à base hybride dans l'application de visualisation.

3. Le procédé selon la revendication 1, dans lequel l'utilisation séparée de la pose basée sur la vision et de la pose basée sur des capteurs dans l'application de visualisation comprend :
l'invitation faite à un utilisateur à déplacer le dispositif mobile (100) dans une direction spécifique avec le premier graphisme et le deuxième graphisme, dans lequel le déplacement du dispositif mobile (100) dans la direction spécifique accélère la convergence de l'étalonnage des capteurs inertiels (110).

4. Un dispositif mobile (100) comprenant :

un moyen pour déterminer une pose du dispositif mobile (100) basée sur la vision, en utilisant des images capturées par une caméra (108) du dispositif mobile (100), en extrayant une ou plusieurs caractéristiques, et en comparant les caractéristiques extraites à des caractéristiques provenant d'une image de référence, dans lequel l'image de référence est une image précédente ou un modèle d'une cible connue ou apprise ;
un moyen pour déterminer une pose du dispositif mobile (100)basée sur des capteurs, en utilisant des capteurs inertiels (110) du dispositif mobile (100), dans lequel la pose basée sur la vision est une première estimation d'une position et d'une orientation du dispositif mobile (100) qui est déterminée en utilisant des techniques de vision par calculateur, et la pose basée sur des capteurs est une deuxième estimation de la position et de l'orientation du dispositif mobile (100) qui est déterminée en utilisant des capteurs inertiels (110) du dispositif mobile (100) ; et
un moyen pour utiliser séparément la pose basée sur la vision et la pose basée sur des capteurs dans une application de visualisation, dans lequel le moyen pour utiliser séparément la pose basée sur la vision et la pose basée sur des capteurs dans l'application de visualisation affiche un premier graphisme basé sur la pose basée sur la vision et un deuxième graphisme basé sur la pose basée sur des capteurs,
dans lequel l'utilisation séparée de la pose ba-

sée sur la vision et de la pose basée sur des capteurs dans l'application de visualisation comprend la réalisation d'un étalonnage des capteurs inertiels (110), dans lequel un utilisateur est invité à déplacer le dispositif mobile (100) le long de trajectoires pour l'étalonnage sur la base du premier graphisme et du deuxième graphisme pour collecter des données qui sont introduites dans un algorithme d'étalonnage.

5. Le dispositif mobile (100) selon la revendication 4, comprenant en outre un moyen pour déterminer une pose à base hybride du dispositif mobile (100) sur la base d'une combinaison de la pose basée sur la vision et de la pose basée sur des capteurs, ou d'une combinaison de données provenant des capteurs inertiels (110) et de données provenant de la caméra (108), et dans lequel le moyen pour utiliser séparément la pose basée sur la vision et la pose basée sur des capteurs dans l'application de visualisation utilise en outre la pose à base hybride.

6. Le dispositif mobile (100) selon la revendication 4, dans lequel le moyen pour utiliser séparément la pose basée sur la vision et la pose basée sur des capteurs dans l'application de visualisation invite un utilisateur à déplacer le dispositif mobile (100) dans une direction spécifique avec le premier graphisme et le deuxième graphisme, dans lequel un déplacement du dispositif mobile (100) dans la direction spécifique accélère la convergence de l'étalonnage des capteurs inertiels (110).

7. Un support de stockage comportant du code de programme stocké dessus, comprenant :
du code de programme qui, quand il est exécuté sur un processeur d'un dispositif mobile (100) équipé d'une caméra (108) et d'un ou plusieurs capteurs inertiels (110), amène le processeur à réaliser le procédé selon une des revendications 1 à 3.

8. Un programme de calculateur comprenant du code de programme qui, quand il est exécuté sur un processeur d'un dispositif mobile (100) équipé d'une caméra (108) et d'un ou plusieurs capteurs inertiels (110), amène le processeur à réaliser le procédé selon une des revendications 1 à 3.

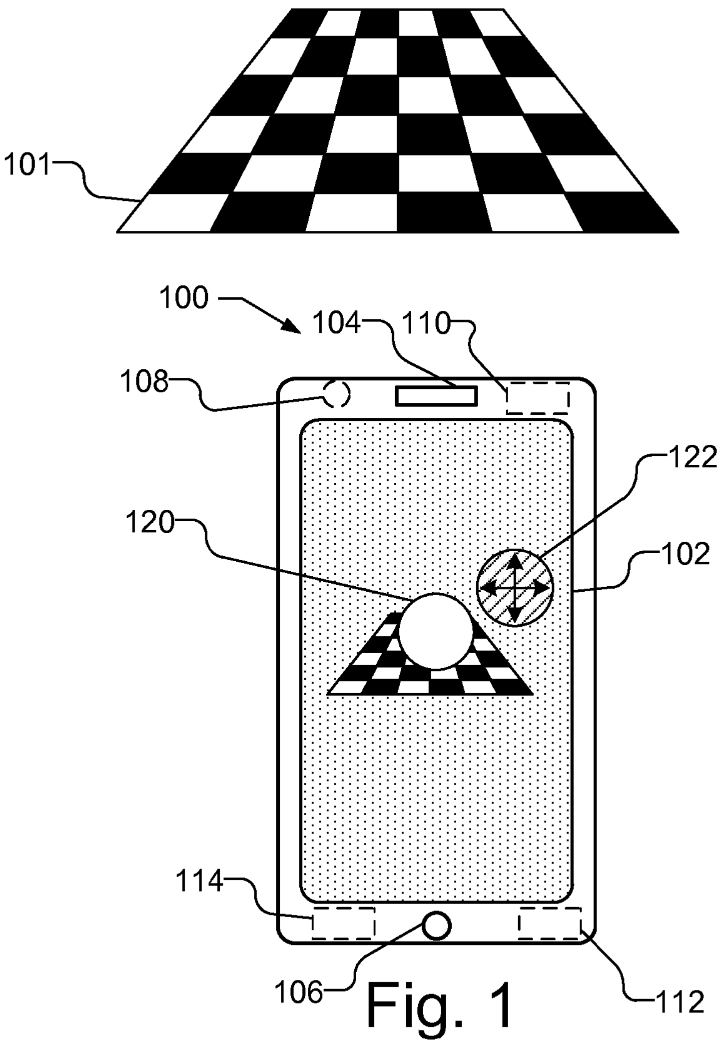
101
100
104
110
108
122
120
102
114
106
112

Fig. 1

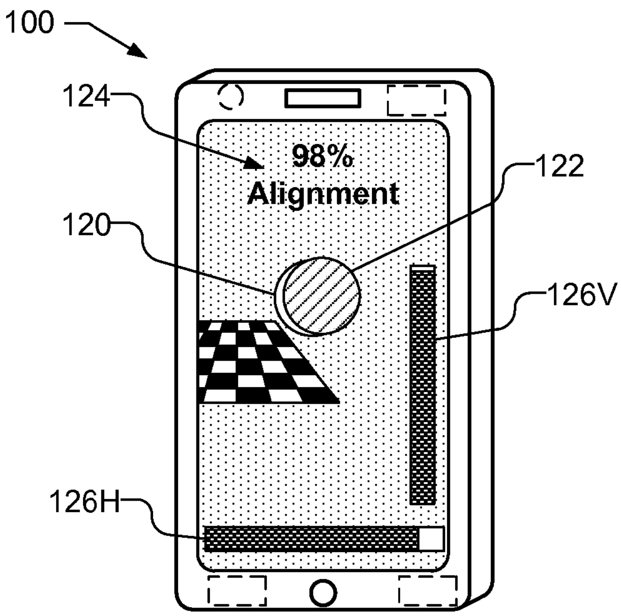
100
124
98%
Alignment
122
120
126V
126H

Fig. 2

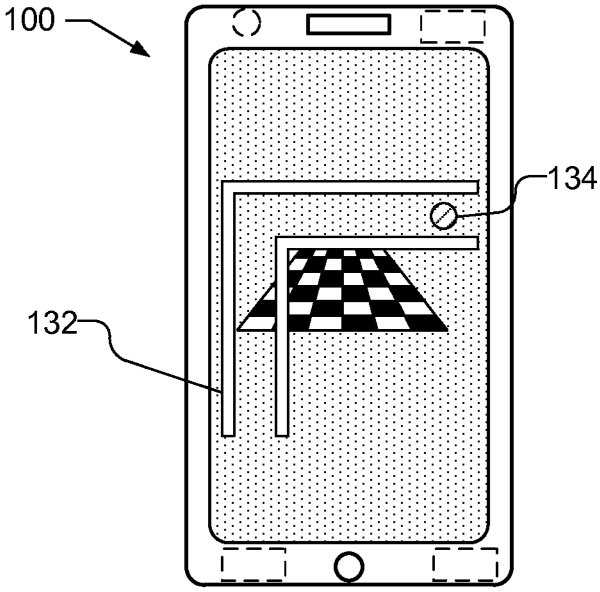

Fig. 3

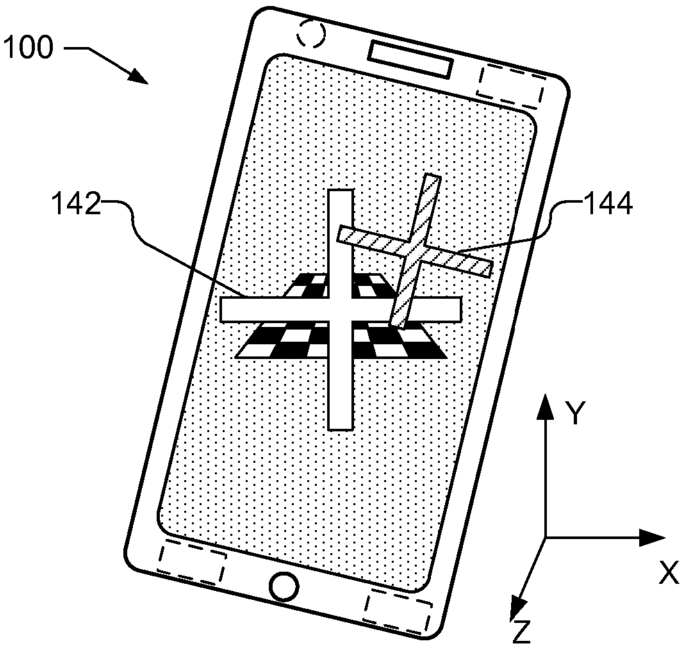

Fig. 4

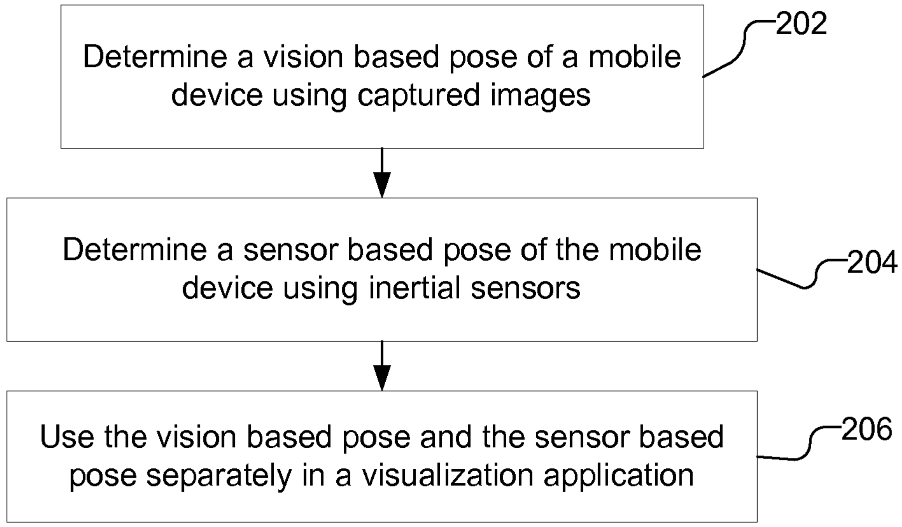
Determine a vision based pose of a mobile device using captured images
202
Determine a sensor based pose of the mobile device using inertial sensors
204
Use the vision based pose and the sensor based pose separately in a visualization application
206

Fig. 5

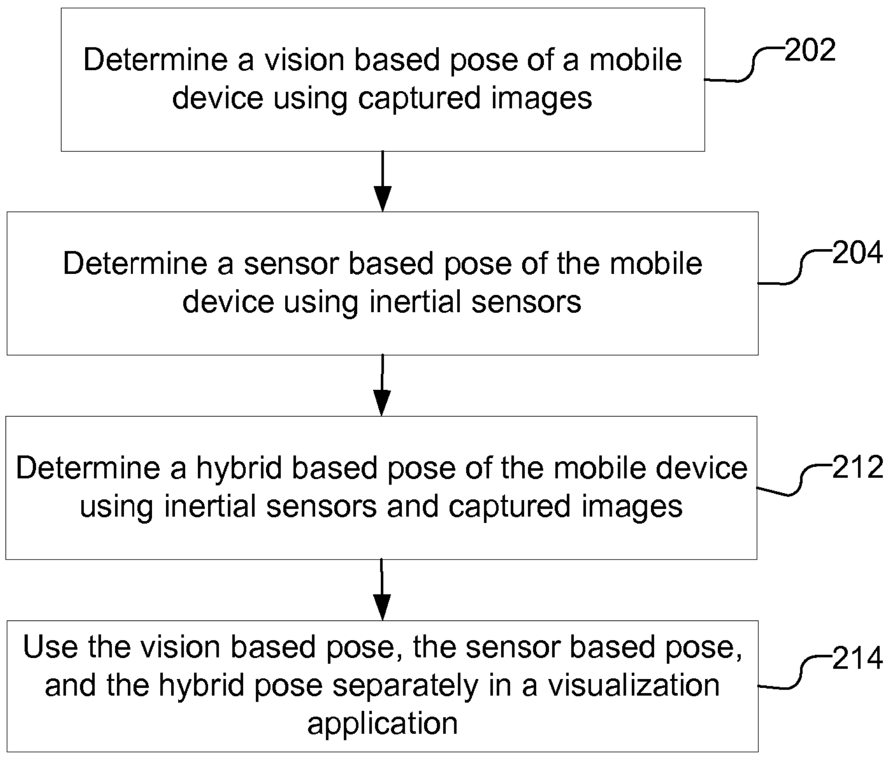
Determine a vision based pose of a mobile device using captured images
202
Determine a sensor based pose of the mobile device using inertial sensors
204
Determine a hybrid based pose of the mobile device using inertial sensors and captured images
212
Use the vision based pose, the sensor based pose, and the hybrid pose separately in a visualization application
214

Fig. 6

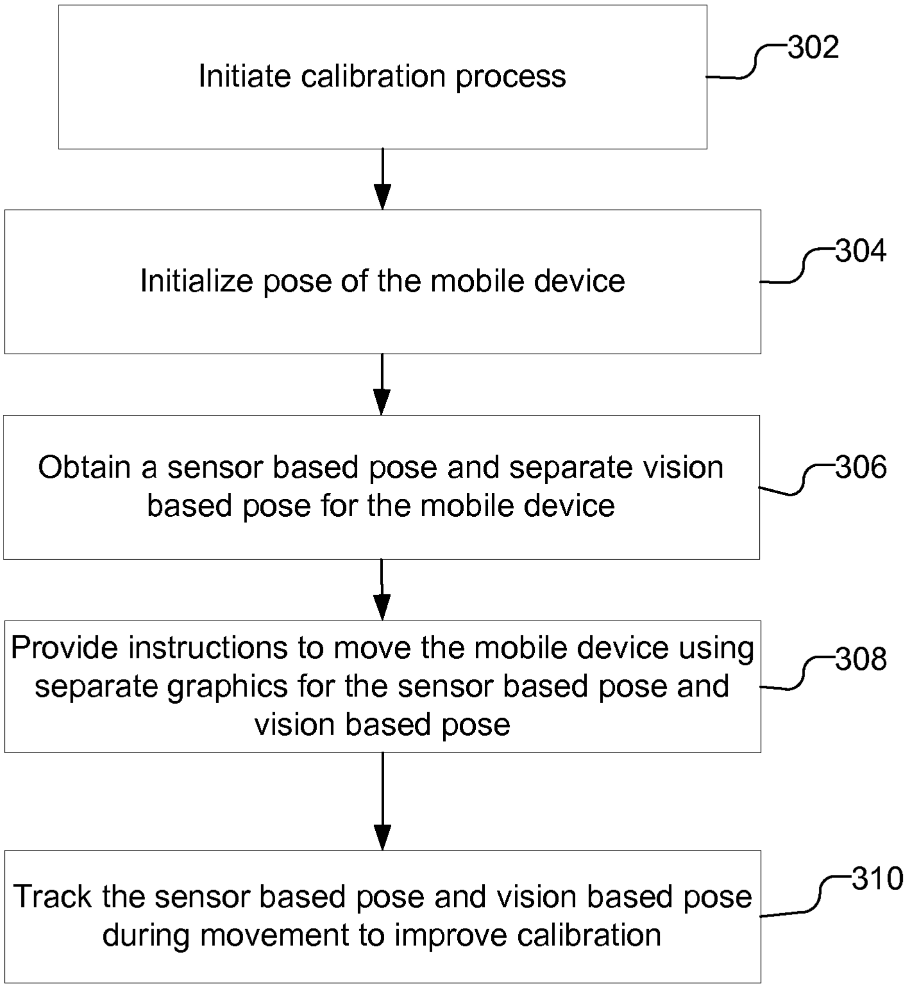
Initiate calibration process
302
Initialize pose of the mobile device
304
Obtain a sensor based pose and separate vision based pose for the mobile device
306
Provide instructions to move the mobile device using separate graphics for the sensor based pose and vision based pose
308
Track the sensor based pose and vision based pose during movement to improve calibration
310

Fig. 7

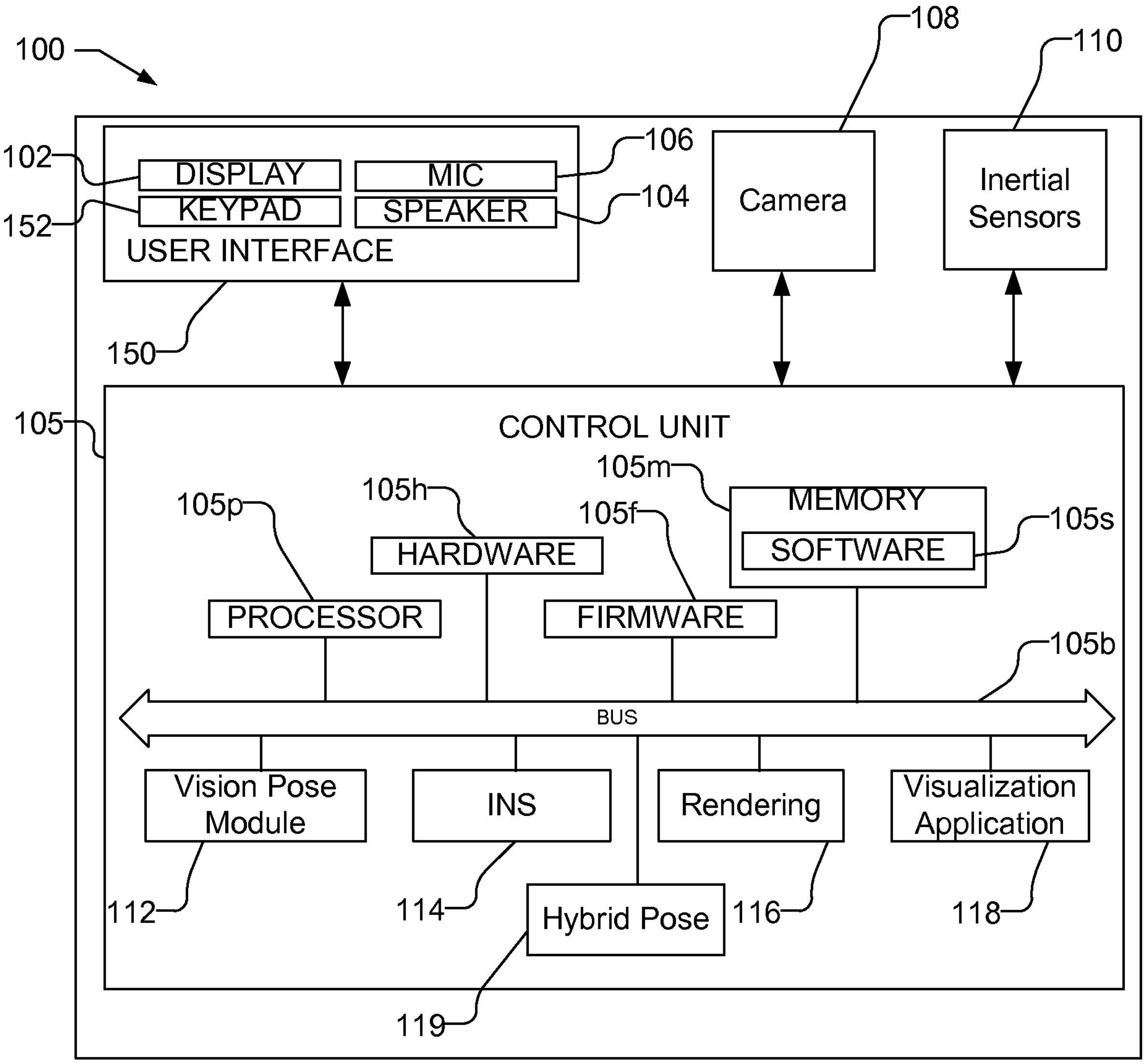
100
102
152
DISPLAY
KEYPAD
MIC
SPEAKER
106
104
USER INTERFACE
150
108
Camera
110
Inertial Sensors
105
CONTROL UNIT
105m
MEMORY
SOFTWARE
105s
105p
105h
HARDWARE
105f
PROCESSOR
FIRMWARE
105b
BUS
Vision Pose Module
INS
Rendering
Visualization Application
112
114
Hybrid Pose
116
118
119

Fig. 8

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- EP 1708139 A2 **[0004]**
- EP 1501051 A2 **[0005]**